INVESTIGATION OF THE ELECTROSTATIC INTERACTIONS OF THE T7 REPLISOME

A thesis presented to the faculty of the Graduate School of Western Carolina University in partial fulfillment of the requirements for the degree of Master of Science in Chemistry.

By

Brittni M. Foster

Advisor: Dr. Jamie R. Wallen
Assistant Professor
Department of Chemistry & Physics

Committee Members: Dr. Indrani Bose, Biology
Dr. Jack Summers, Chemistry & Physics

June 2016

ACKNOWLEDGMENTS

First, I would like to thank my thesis advisor Dr. Jamie Wallen for opening my eyes to the field of biochemistry. He encouraged me to continue my studies and pursue a PhD at Wake Forest University in the fall. I would like to thank my committee members Dr. Jack Summers and Dr. Indrani Bose for being available to answer questions and give advice toward my project. Thank you to Dr. Maria Gainey for being the local virologist expert on optimizing my T7 phage experiments. A special thank you goes out to Lynley Hardie and Wesley Bintz; these two were always available to help in anyway possible during my time at Western Carolina University. A thank you to Jenny Collins and Maggie Carver, two other students working in Dr. Wallen's lab that assisted on this project. Randal Pinkston thank you for helping fix the Äkta when it would not connect. Lastly, I would like to thank my family for always encouraging and supporting me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

gp5 ..... polymerase, Gene product 5 of the T7 phage

Pol ..... polymerase

gp4 ..... helicase-primase, gene product 4 of the T7 phage

Trx ..... thioredoxin

gp2.5 ..... SSB, gene product 2.5 of the T7 phage

SSB ..... Single Stranded Binding Protein

dNTP ..... deoxynucleotide triphosphate

ssDNA ..... single-stranded DNA

dsDNA ..... double stranded DNA

Fbp ..... front basic patch

TBDbp ..... thioredoxin binding domain basic patch

SAXS ..... small angle X-ray scattering

ABSTRACT

INVESTIGATION OF THE ELECTROSTATIC INTERACTIONS OF THE T7 REPLISOME

Brittni M. Foster, Masters of Science in Chemistry

Western Carolina University (June 2016)

Advisor: Dr. Jamie R. Wallen

DNA replication is a process in which genetic information is copied and passed on from generation to generation. Multiple proteins of specialized function must transiently assemble and communicate to rapidly and accurately copy DNA. We study replication in bacteriophage T7, a virus that infects Escherichia coli (*E. coli*), because it is a relatively simple model system that recapitulates all essential activities seen in more complex systems in just four proteins. Recently our laboratory determined a crystal structure of an electrostatic interaction between T7 primase-helicase and DNA polymerase, and the structure revealed the acidic C-terminal tail of the T7 primase-helicase binds to a basic patch on the surface of the polymerase. The C-terminal tail of the primase-helicase contains a phenylalanine residue that is essential for binding polymerase, and our structure reveals that this phenylalanine binds to a hydrophobic cleft near the basic patch. Our crystal structure also unexpectedly revealed an interaction between a nucleotide triphosphate and tryptophan 160 located in the exonuclease active site, and this interaction has shed light on how the exonuclease domain of T7 polymerase recognizes a DNA substrate. To validate our crystal structure, point mutations were generated along the basic patch of the polymerase as well as the hydrophobic pocket to disrupt the observed interactions.

Using in vivo methods such as phage complementation and multi-step growth curve we are able to directly test the effects on the virus replication during infection of *E. coli*. Phage complementation results demonstrate that these mutations F487A, I569A, R687A and R590A show a 3-log and 4-log difference in growth compared to WT. Plaque assay results show that R590A is critical for replication, as the virus shows no growth under identical conditions for WT. The other mutations I569A, I569A, and F487A were also impaired during the plaque assay and similar to R590A showed no plaque formation. R687A showed substantially smaller plaques compared to WT and we were unable to calculate a titer. No plaque formation and smaller plaques indicate that the virus is heavily impaired in DNA replication. Under identical conditions to WT, R590A show no growth over a 15 hour period where as WT reaches a maximum amount of virus and plateaus at 4 hours. Preliminary data for I569A and F487A show these mutations also result in impaired and reach a lower maximum viral concentration compared to WT. In the crystal structure R590 of the polymerase makes a critical contact with the C-terminal tail, and in agreement with the structure, mutation of R590 to alanine suggests that this residue serves an essential function in DNA replication.

Phage complementation results for the two mutations in the exonuclease active site, W160A and 5A7A, gave similar results compared to WT at a log of $1x10^{12}$ PFU/mL. The double mutation D5A and E7A (5A7A) is unable to coordinate Mg^{+2} causing exonuclease activity to be diminished. Results for rolling circle reveal that polymerase with a 5A7A mutation can perform strand displacement synthesis without helicase. Future work will focus on continuing Rolling circle, growth curve, and primer extension assays.

CHAPTER 1 INTRODUCTION

1.1 MOTIVATION

DNA replication is a complex process requiring a coordinated system consisting of a multitude of proteins. A simple organism such as the bacteria *Escherichia coli*(*E. coli*) needs 12 proteins to replicate their DNA while 40-50 human proteins are needed to accurately copy both strands.[3,4] Multiple proteins of specialized function must transiently assemble and communicate to rapidly and accurately copy DNA of all genomes. These proteins form a highly coordinated machine called a replisome.[5] If anything goes wrong during DNA replication, mutations can occur, leading to numerous diseases. Due to the complexity of the human replisome, many of the mechanisms for protein-protein interactions in the human system are still unknown. Because of this, simpler replication systems are used as models.[6] Bacteriophage T7, a virus that infects *E. coli*, is a model system that can recapitulate all essential activities of DNA replication with only four proteins.[1]

This project brings understanding of how these four replication proteins interact not only in T7 but also in higher organisms. An understanding of the T7 replisome will shed light on essential interactions in replisomes of pathogenic bacteria. A study of the T7 system can provide a selective way to attack the replication system in pathogenic bacteria without adverse effects to humans. Recently, a crystal structure was determined showing an electrostatic interaction between the T7 primase-helicase C-terminal tail and the T7 DNA polymerase (unpublished data from our lab). For DNA replication to occur these two proteins must interact and the interaction between the C-terminal tail of the primase helicase and DNA polymerase is essential. This project focuses on validating this specific electrostatic interaction between the T7 DNA polymerase (pol) and the C-terminal tails of both the helicase and single-stranded DNA binding proteins (SSB). Since the interaction between the SSB and the

pol is seen in both T7 and pathogenic bacterial DNA replication but is absent in humans, this interaction could be targeted by therapeutics.[7–9]

1.1.1 DNA REPLICATION

To successfully replicate duplex DNA several events need to occur. The double stranded DNA (dsDNA) must be unwound, primer synthesis needs occur to initiate lagging strand synthesis, the antiparallel strands need to be identically copied, and an editing method needs to be incorporated to make sure the correct nucleotides have been put in place. Figure 1 illustrates the synthesis of duplex DNA for the leading and lagging strands of DNA.[5]

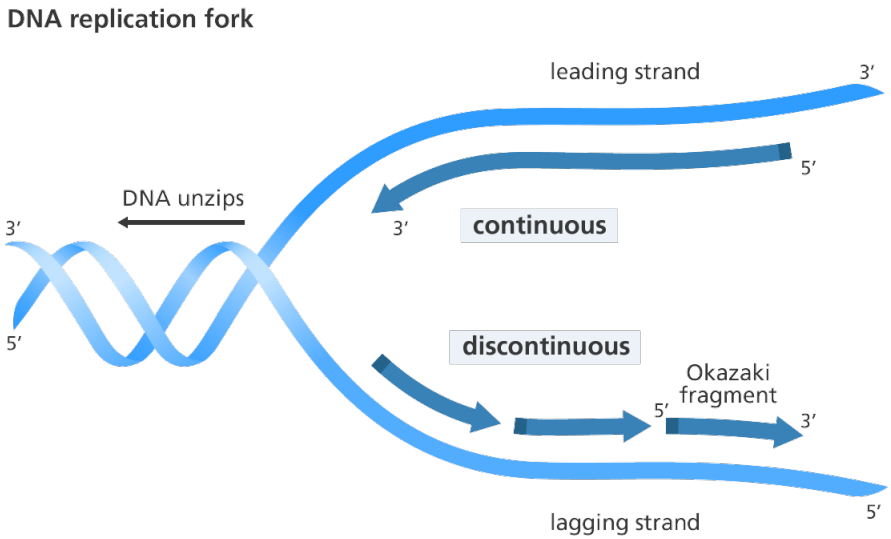

Figure 1. An illustration of leading and lagging strand synthesis of duplex DNA.

Double stranded DNA is antiparallel in nature, and replisomes can only synthesize DNA in one direction. DNA is unwound by a motor protein called a helicase. For T7 the C-terminal helicase domain of the primase-helicase enzyme will unwind the DNA in the 5'→ 3' direction. As the DNA is being unwound both strands will be copied at the same time by a DNA polymerase. In T7, the DNA polymerase synthesizes DNA in the 3'→ 5' direction.

At least two polymerases are needed in the replisome to synthesize both leading and lagging strands simultaneously. To overcome the antiparallel orientation of DNA the leading strand is synthesized in short stretches (∼3,000 bases) called Okazaki fragments. To create the Okazaki fragments the lagging strand needs to align its polarity with the leading strand causing a replication loop to form.[1] The lagging strand polymerase begins replication of the discontinuous Okazaki fragments using RNA primers synthesized by the N-terminal primer domain of the primase-helicase. The small fragments will later be joined by a DNA ligase. The T7 polymerase has 3'→ 5' exonuclease activity, making it able to check that the correct nucleotides were put into place during replication.

1.2 T7 BACTERIOPHAGE

T7 is a lytic phage meaning that it destroys the host once it has sufficiently replicated its genome and assembled new viral particles.[10] T7 attaches to the host cell and injects its 40 kilobase (kb) genome that encodes approximately 50 proteins into the bacterial host.[11] This phage replicates its DNA using the four essential proteins, three of which are encoded from its own genome and one that is hijacked from the host. These four proteins are coordinated and act as one machine called a replisome.[12] Each protein in the replisome has an essential function; if any one of these proteins are missing, DNA will not be synthesized. Figure 2 shows the current model for how the four T7 proteins assemble to copy DNA. The first protein is the primase-helicase (purple circle). The N-terminal domain catalyzes RNA primer synthesis needed to initiate DNA synthesis of the lagging strand, while the C-terminal helicase domain catalyzes nucleotide-dependent DNA unwinding. The second protein is the DNA polymerase (green circle) that copies the DNA. Thioredoxin or Trx (yellow sphere attached to the polymerase) binds to polymerase to aid in copying the DNA by increasing its processivity (number of nucleotides that are copied in a single binding event). Trx is the only protein not encoded in the T7 genome and is hijacked from the host *E. coli*.[1] Poly-

merase and thioredoxin are two separate proteins and they bind together to form a high affinity complex (pol/trx) with a Kd of ∼5nM.[13] The fourth protein is the single-stranded DNA-binding protein (SSB, Orange circle). SSB binds and protects single stranded DNA as well as interacts with both the helicase and polymerase to coordinate the replisome into a single functional unit.[1] It is important to note that both the primase-helicase and the SSB have an acidic C-terminal tail that is represented by the red lines in Figure 2.

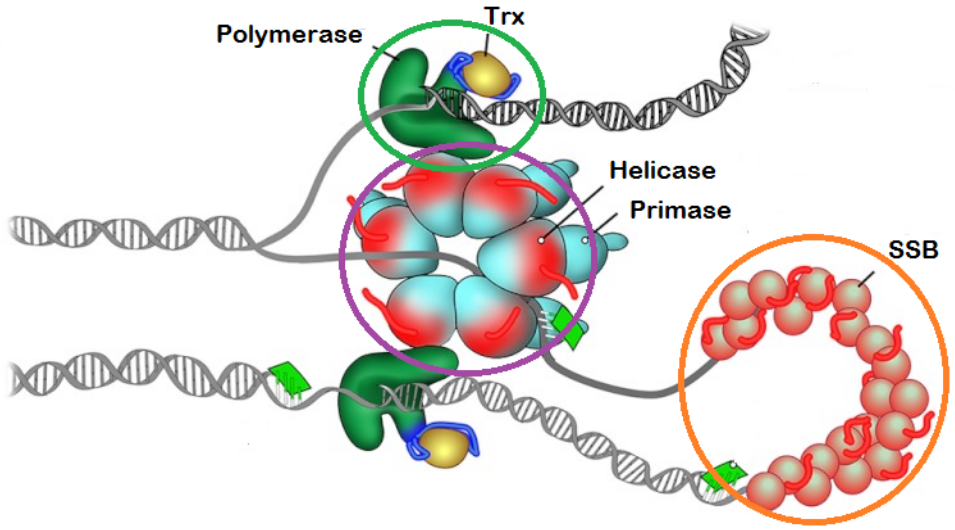

Figure 2. [1] The current model of the four T7 proteins working together to replicate DNA. Circled in green is the polymerase, circled in orange is the single-stranded DNA binding protein, circled in purple is the primase-helicase.

1.2.1 MODES OF REPLICATION

When replicating DNA, the primase-helicase and the polymerase have at least three modes of binding known as the polymerization mode, the priming mode, and the electrostatic mode.[1] The electrostatic mode is the only mode of binding that requires the electrostatic interaction involving the acidic C-terminal tails of the helicase and SSB, and this is the mode that is the focus of this project. The electrostatic interaction involves the highly acidic C-terminal tails of helicase and SSB with the two basic patches of the polymerase (Figure 4).[2]

To initiate synthesis the helicase binds to the front basic patch(Fbp) of the polymerase to aid in loading it onto DNA. In the event that pol dissociates from DNA The acidic C-terminal tail of the helicase is important for keeping the dissociated polymerase tethered to the moving replisome. In the current model, the helicase C-terminal tail will bind to the thioredoxin binding domain basic patch (TBDbp) when the pol dissociates from the DNA and the pol will be rotated so the helicase tail can interact with the FBP to reload onto DNA. Figure 3 demonstrates how the helicase interacts with both the TBDbp and the FBP.[2,14]

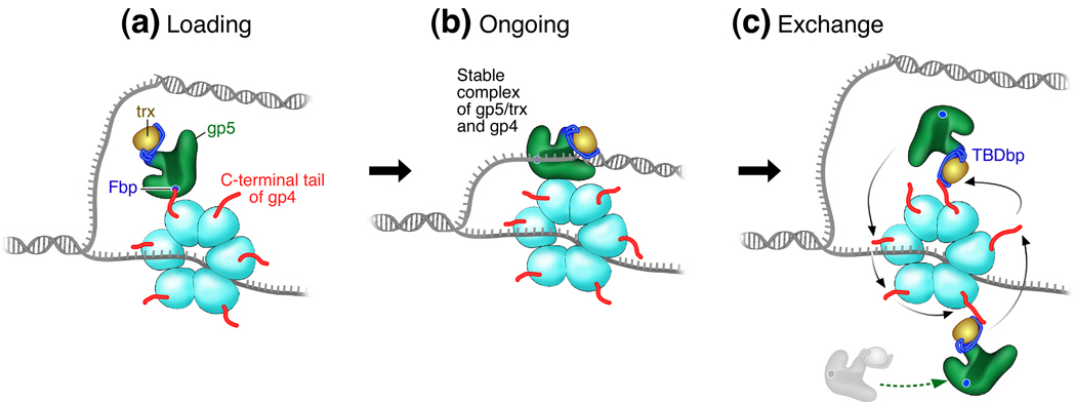

Figure 3. The helicase must interact with the polymerase during DNA synthesis. The FBP and the TBDbp are important for keeping the polymerase tethered to a moving replisome. In the current model to begin replication the C-terminal tail of the helicase will bind to the Fbp of the pol and load it onto the DNA as seen in figure A and B. In the event that the polymerase dissociates from DNA the C-terminal tail of the helicase will bind to the TBDbp and rotate the polymerase to bind to the Fbp to load the polymerase back onto DNA.[2]

Figure 4 illustrates the location of the TBD and the Fbp on the polymerase.[2] This electrostatic interaction is essential for increasing the processivity of the replisome.[2] The T7 genome consists of approximately 40kb which needs to be replicated for phage survival. When the pol/trx is able to interact with the helicase the replisome has a processivity of 17kb. If the acidic C-terminal tail of the helicase is deleted, the processivity of strand displacement synthesis decreases from 17kb to 5kb, indicating that the electrostatic interaction is needed for high processivity between the helicase and the polymerase.[1,2] Strand displacement synthesis

is the ability of an enzyme to displace DNA encountered essentially unwinding dsDNA and synthesizing a new strand.

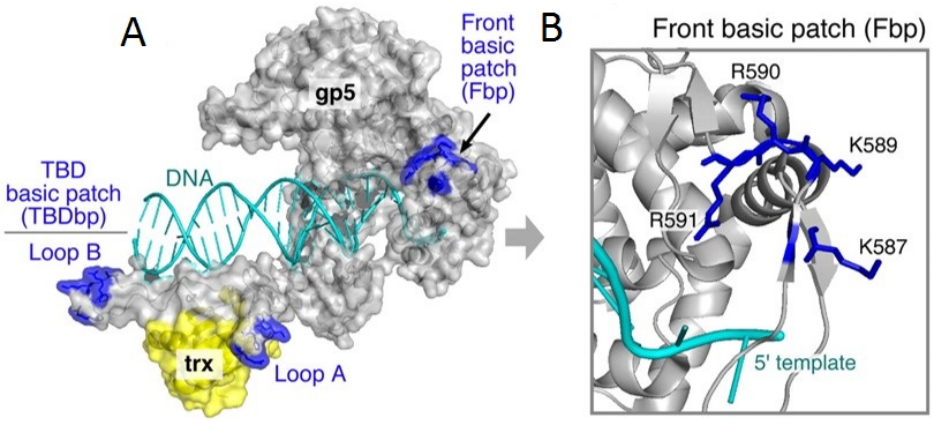

Figure 4. Illustration representing the polymerase/thioredoxin complex structure showing the locations of the front basic patch and the thioredoxin basic patch[2].

1.3 CRYSTAL STRUCTURE

Recently, a three-dimensional structure of the C-terminal tail of T7 primase-helicase bound to the polymerase was determined by Dr. Wallen using X-ray crystallography (Figure 5). The crystal structure shows that the C-terminal tail binds in a pocket near Fbp, and the very C-terminal phenylalanine, which has been shown to be essential for the interaction, is buried in the hydrophobic cleft of the polymerase (Figure 5). The C-terminal tail of the helicase is essential. If it is deleted, neutralized, or the C-terminal phenylalanine is mutated, replication no longer occurs.[2,15,16] As seen in Figure 6, the C-terminal tail binds the Fbp such that the aromatic phenylalanine lies in the hydrophobic pocket, and the tail is further anchored by two salt bridges involving the D565 side chain and F566 carboxylate of the helicase with R687 and R590 of the polymerase, respectively. K587, K589, R590, and R591 (Figure 4B) are among the basic residues that have been previously neutralized in the Fbp and are known to be vital for the interaction between the primase-helicase acidic C-terminal

tail and the polymerase.[2] R590 is the amino acid seen anchoring the helicase tail to the front basic patch in the crystal structure. We hypothesize that if we mutate R590 the interaction between the helicase and the polymerase will weaken and T7 will not be able to replicate its DNA efficiently.

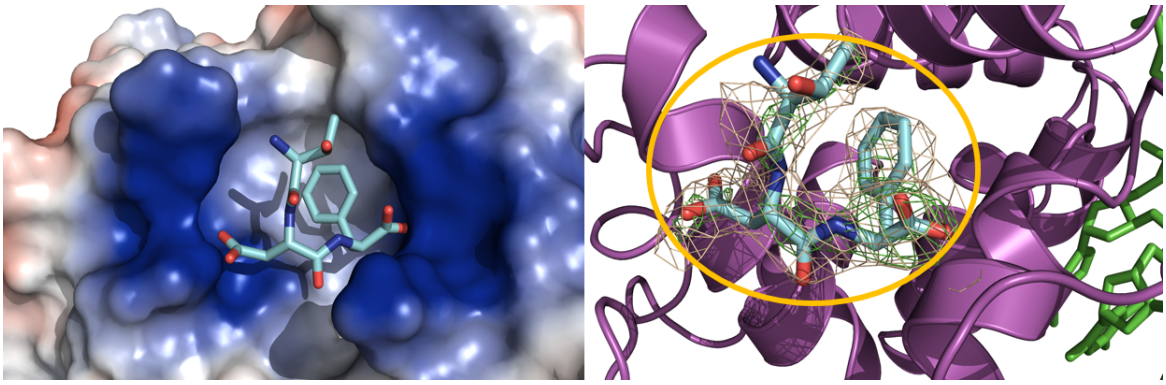

Figure 5. Crystal structure showing T7 helicase peptide (yellow circle in right hand panel) bound to the polymerases hydrophobic pocket. The panel on the left is color coded based on charge, blue represent basic amino acids while white represents neutral amino acids.

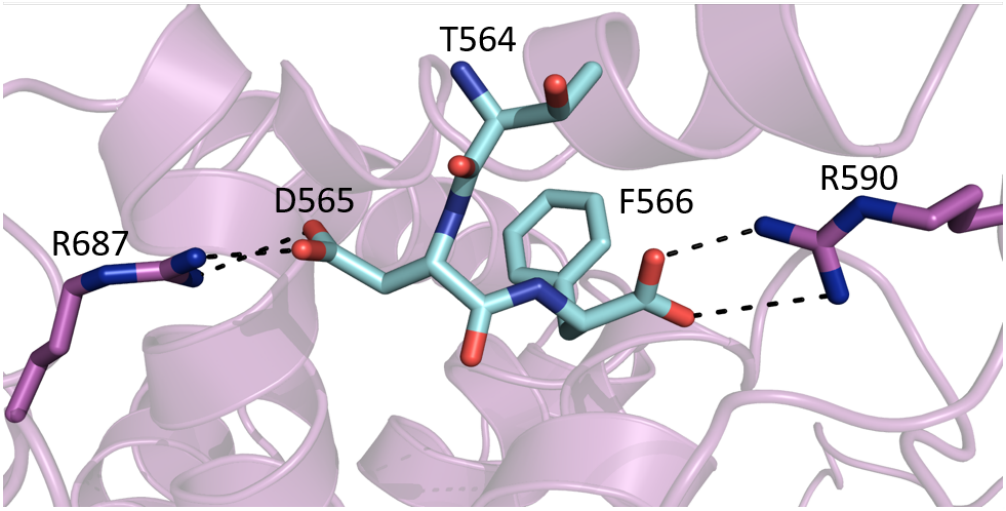

Figure 6. Close up of the crystal structure illustrating the salt bridges being formed between the peptide and the polymerase in the front basic patch.

The C-terminal phenylalanine of the helicase fits perfectly into the polymerase hydropho-

bic pocket, like a lock and key. The amino acids that make up the hydrophobic pocket are F487, I569, L163, I593, T572. Mutating any one of the amino acids that make up the hydrophobic pocket to an alanine we hypothesize will change the shape decreasing or diminishing this hydrophobic interaction. For this study we have preformed site directed mutagenesis on the following amino acids L613, I569, and F487, mutating each separately to an alanine. Figure 7 shows the surface representation of the peptide binding in the hydrophobic pocket.

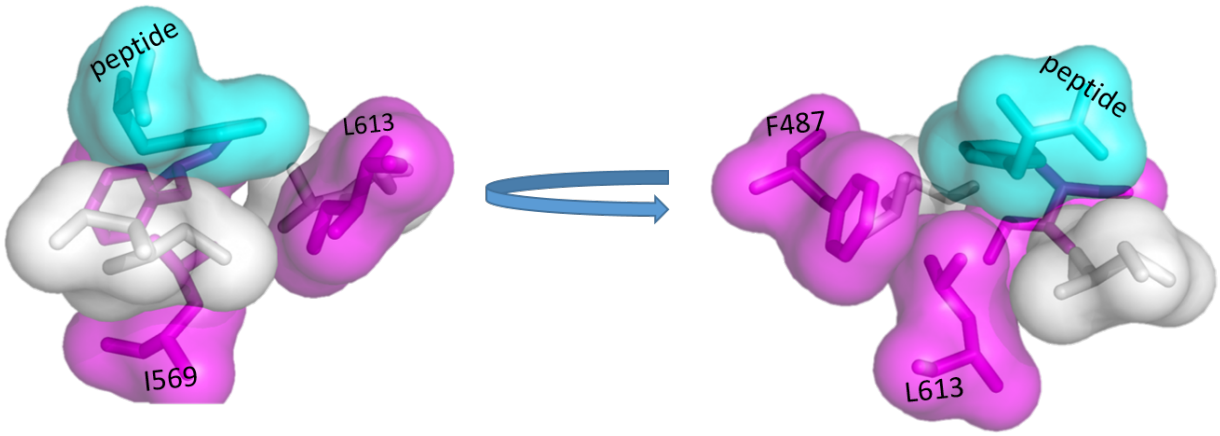

Figure 7. View of the amino acids surrounding the helicase tail peptide. The amino acids shown in purple are the amino acids that have been mutated to an alanine for this project. The amino acids shown in white have yet to be mutated.

1.3.1 EXONUCLEASE

Replicative DNA polymerases, such as T7 DNA polymerase, are highly accurate motors. They require a correct nucleotide to be added to the primer template in order to move forward. For every 10^3-10^5 nucleotides incorporated, $\sim$1 mistake will occur.[1] If a mistake does occur the T7 DNA polymerase contains a exonuclease site that will remove the incorrect nucleotide. Having the ability to proof-read increases pols fidelity by $\sim$100 fold. The exonuclease can remove the incorrect nucleotide on ssDNA in the 3'$\rightarrow$ 5' direction. When a mistake is made the primer template needs to be shuttled over 30$\mathring{A}$ to the exonuclease site where it will be removed. Until recently there has been difficulty crystallizing DNA in the

exonuclease site for the T7 DNA polymerase.[1]

In the same crystal structure with peptide bound (described above), a nucleotide (dTTP) is seen in the exonuclease site. In a second crystal structure of polymerase 5A7A bound to helicase in the presence of dT15 ssDNA, ssDNA was also found in the exonuclease active site. There has never before been a crystal structure of the T7 polymerase showing ssDNA or a nucleotide in the exonuclease site. It is not known how the DNA is shuttled to the exonuclease site to remove the mistake for T7 DNA pol. The crystal structures in Figure 8 reveals the W160 aromatic side chain stacking against a nucleotide in each crystal structure.

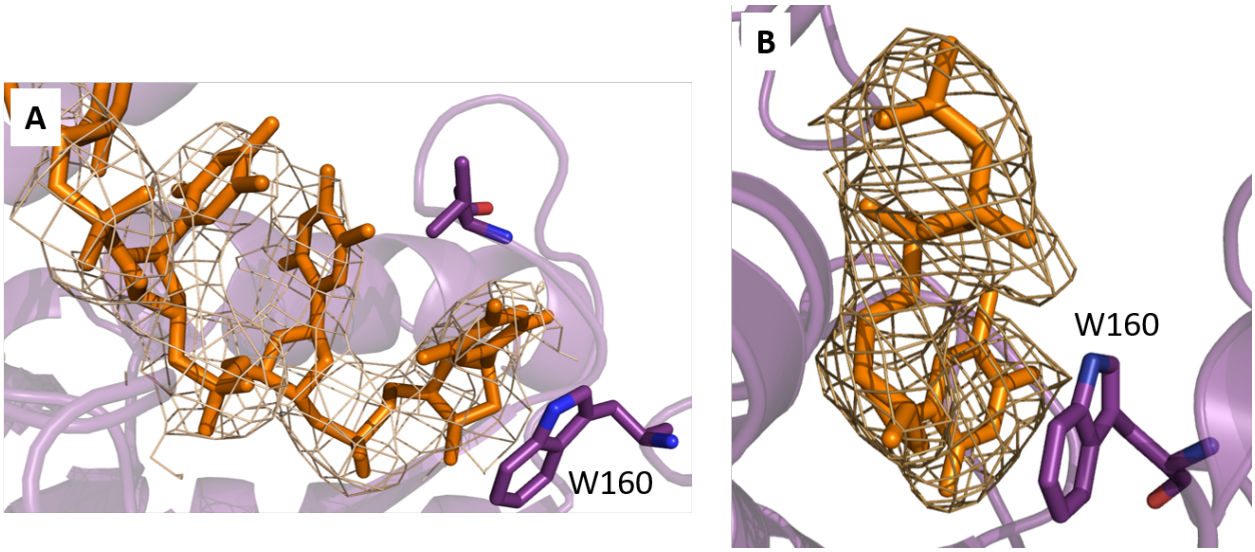

Figure 8. Two crystal structures showing the electron density of a nucleotide interacting with W160 in the exonuclease site. Panel A on the left shows a crystal structure with ssDNA in the exo site and panel B shows a single nucleotide (dTTP) in the exo site. The base of the nucleotide is seen stacking with the aromatic side chain of W160.

In the exonuclease site a sequence alignment seen in Figure 9 shows that residues D5 and E7 are conserved in all 63 T7-like polymerases (30% or greater identity). These residues coordinate Mg^{+2} that facilitates the exonuclease attack on DNA. Along with D5 and 7E W160 is an aromatic residue in all sequences, as either a W or an F. With the sequence alignment and the crystal structure we hypothesize that this W160 is an essential amino acid for stabilizing the DNA in the exonuclease site.

Figure 9. Above is a logo graph showing the alignment of T7 Pol with 63 T7-like polymerases at 30% or greater identity. Residues 5 and 7 coordinate Mg^{+2} and are absolutely conserved in all Pols. W160 is an aromatic in all sequences, either as a W or an F.

If W160 is an essential amino acid in the exonuclease site then mutation of this residue should alter exonuclease activity. W160 has previously been mutated to an A in our lab (unpublished). To determine if the W160A pol could replicate a normal primer template and a mismatch primer template like WT, a sequencing gel was ran(Figure 10). The W160A pol when loaded onto a primer template containing no mutation is able to replicate the template similar to WT pol. Unlike WT the W160A was not able to replicate the mismatch primer template. WT pol must remove the mutation before replication can begin. This preliminary data suggests that the W160 is an essential amino acid for a functioning exonuclease.

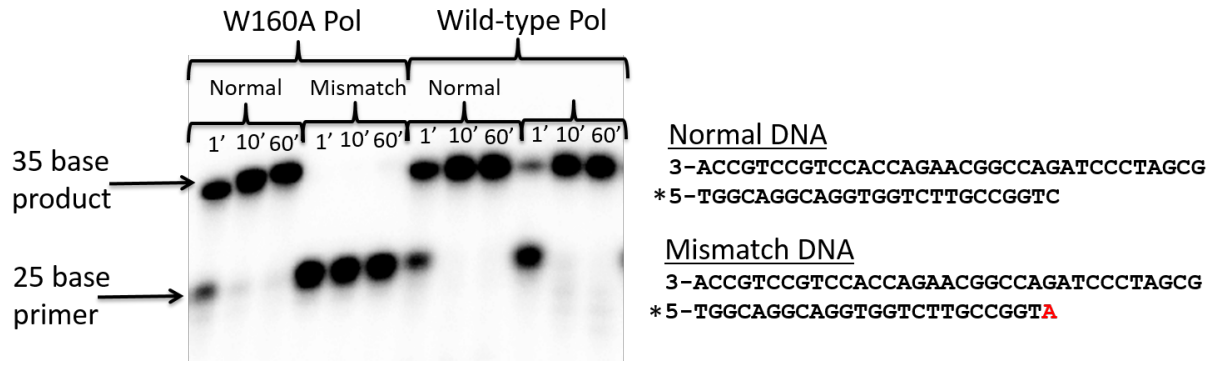

Figure 10. Sequencing gel of W160A and WT pol. This assay compares the w160A ability to extend DNA on both a normal DNA primer template and a mismatch primer template. WT pol can extend DNA on both the normal and mismatch primer template. The W160A can extend DNA on normal DNA primer template but can not complete DNA synthesis on primer template containing the mismatch. This demonstrates the importance of the residue W160 for the exonuclease activity of the T7 DNA polymerase. If the polymerase can not correct a mistake in this case a mismatch then the T7 DNA pol can not continue synthesizing DNA. (Data obtained previously by Dr. Wallen)

This thesis will focus on 1) validating the electrostatic interaction between the C-terminal tail of the helicase and the Fbp of the pol. Understanding at the molecular level how this interaction coordinates the replisome and 2) Verify that W160 is an essential amino acid in stabilizing the DNA in the exonuclease site.

1.4 SSB AND PATHOGENIC BACTERIA

The SSB C-terminal tail is very similar to that of the primase-helicase. Both tails have many acidic residues essential for interacting with the polymerase, as seen in Figure 11. The last two amino acids are identical which includes the imperative C-terminal phenylalanine and many acidic residues are conserved. The acidic residues of the tail along with the C-terminal phenylalanine are essential for DNA replication.[17–19] With the similarities between the SSB and helicase C-terminal tail we hypothesize that the SSB tail will bind to the same site of the polymerase Fbp as seen with the helicase peptide in the crystal structure.

Figure 11. Sequence alignment of the helicase C-terminal tail and the SSB C-terminal tail. Notice that both the helicase and the SSB contain the end terminal phenylalanine (denoted by the letter F) and many acidic residues highlighted in red.

It is known that the C-terminal region of the SSB is essential for DNA replication.[1,20,21] The binding of the SSB to polymerase is the same within pathogenic bacteria but not in eukaryotes, making this a perfect interaction to study for drug development.[1] The SSBs from many bacteria are structurally similar to that of T7 and are expected to bind to polymerase in a similar fashion. To illustrate, Figure 12 provides amino acid sequences for the C-terminal tail of T7 SSB protein and several pathogenic bacteria. Note that the C-terminal sequence for all the SSBs in Figure 12 end with the same amino acid, phenylalanine (given the letter F) and that there are many other conserved acidic amino acids in their sequences. Also note that the bacteria in Figure 12 are pathogenic and cause the following diseases: *Staphylococcus* (Staph), *Streptococcus* (Strep), *Salmonellosis* (Salmonella), and *Bordetella Pertussis* (Whooping cough), which are all becoming dangerously resistant to antibiotics. Thus, understanding how SSB binds polymerase in our T7 model system may lead to new drugs that will prevent this protein/protein interaction and kill these pathogenic bacteria.

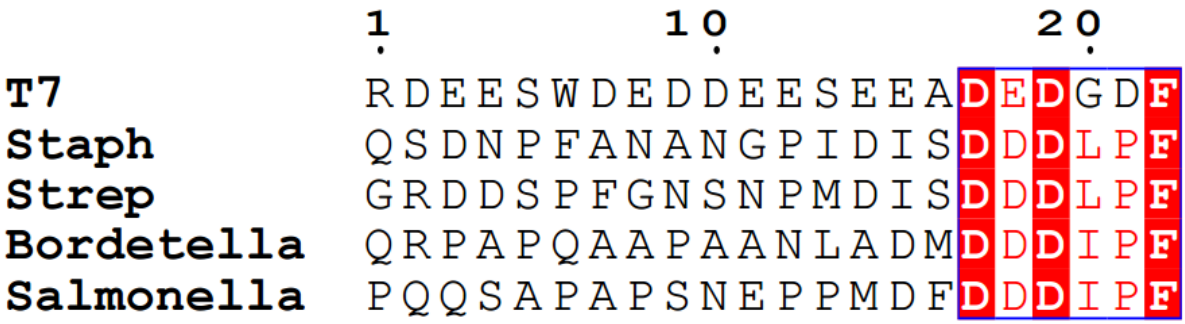

Figure 12. Bioinformatics comparison of T7 SSB tail to bacterial SSBs; staph, strep, bordetella, and salmonella. The end terminal phenylalanine (given the letter F) is conserved in all SSB and is known to interact with the DNA polymerase. The amino acids (D and E) are present in all sequences indicating that all these SSB likely interact with polymerase in the same way as T7 SSB.

CHAPTER 2 INTRODUCTION TO RESEARCH

2.1 OBJECTIVES

This thesis will focus on validating the interactions seen in these three dimensional structures to understand how the electrostatic interaction coordinates the replisome at the molecular level. Additionally, we will verify that the W160 is an essential amino acid in stabilizing the DNA in the exonuclease site.

Several experiments have been performed to validate the crystal structure showing the electrostatic interaction between the C-terminal tail of the helicase and the Fbp of the polymerase along with the W160 as an important residue of the exonuclease site. Site directed mutagenesis was carried out on several of the amino acids in the Fbp of the polymerase including L613, F487, R590, R687, and I569. W160 was also mutated to an A to test the importance of the aromatic nature. The effect of these mutations was studied using phage complementation, rolling circle, multi-step growth curve, and primer extension assay.

2.2 PHAGE COMPLEMENTATION

Phage complementation was used to determine how point mutations of the Fbp affect replication of the T7 phage in vivo. A T7 phage that is missing the polymerase gene is used in these experiments (Δgp5). Without the gene for the polymerase the phage cannot replicate and therefore cannot infect its host *E. coli*. By providing the polymerase gene via a separate DNA plasmid the ability for the phage to replicate is restored and results in an infectious phage as seen in Figure 13.

Mutations of the polymerase gene can easily be made on the plasmid DNA and the effects of these mutations can be studied to determine which specific amino acids are vital for phage growth and survival. If an amino acid that is essential for replication is mutated it can lead to no phage growth, no plaques visible, or a change in plaque phenotype.

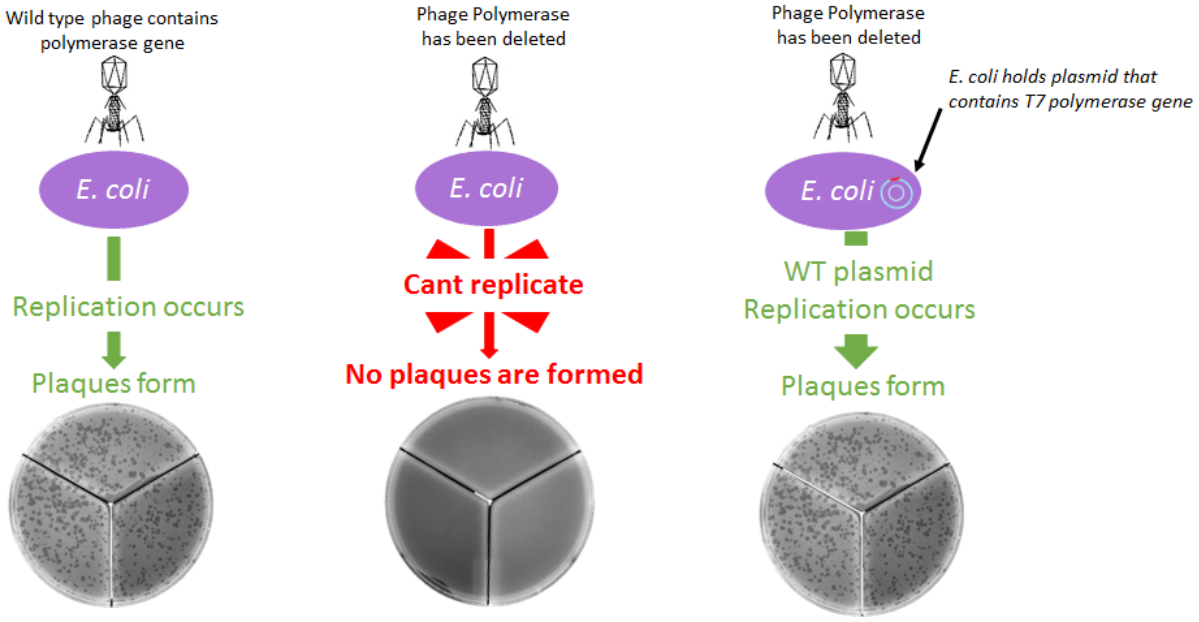

Figure 13. Phage complementation is a technique used to quantify the effects that the point mutations have on the growth of T7 phage in vivo. From left to right Figure A demonstrates wild type plaque formation. The wild type phage contains all genes necessary for the phage survival in *E. coli* producing plaques in the bacterial lawn. Figure B contains a T7 phage with the deleted polymerase gene that is needed for the phage to grow. Without this gene the phage can no longer survive and plaques will not form. Figure C demonstrates how the phage complementation technique is performed. The T7 phage has the polymerase gene deleted(Δgp5) from its genome. A DNA plasmid is transformed into the *E. coli* cells containing the polymerase gene. By adding the polymerase gene to the plasmid it complements the the Δgp5 phage and allows for plaques to once again form. Mutations can easily be made on the plasmid and the effects of phage survival can be quantified.

2.2.1 GROWTH CURVE

A multi-step growth curve is used to study the effects the point mutations have on the phage growth over time. *E. coli* cells are infected with a multiplicity of infection (MOI) of $1x10^{-7}$ to $1x10^{-9}$ of phage. The infected cells are grown at 37°C and samples were taken at 2,4,6 and 15 hours. The samples are spotted onto the bacterial lawn and the plaques are counted to determine a titer of virus. This assay will allow us to compare the mutation's effects on viral growth. If the replisome cannot coordinate due to our mutations, then replication cannot occur and the phage cannot survive. The mutation is detrimental, such that pol and helicase cannot interact, then no viral growth will occur. If the interaction is weakened, then that will become evident by a delayed viral growth over time.

2.3 ROLLING CIRCLE

The rolling circle replication assay is used to test the ability of the polymerase to synthesize DNA and the replisomes ability to coordinate and synthesize DNA. Rolling circle uses a single stranded circular piece of DNA from the virus M13. A primer complimentary to the M13 DNA is bound to a FAM label and is annealed to the ssDNA for polymerase and helicase to load. A dT35 tail is added to the 5' end of the primer. This does not anneal with the M13 DNA and provides a fork-shaped structure that is required for primase-helicase loading. Figure 14 describes how rolling circle is carried out. To test if the mutations alter the ability of the polymerase to synthesize DNA only the pol/trx complex is loaded onto the ssM13. If the pol can synthesize DNA it will be able to complete the circle until it reaches the tail. Without helicase pol/trx cannot perform strand displacement DNA synthesis.[1,2] Once helicase is added strand displacement DNA synthesis can occur and the pol/trx can continuously synthesize DNA around the circle. Under these conditions we are only able to monitor the ability of primase-helicase and polymerase to carry out coordinated leading

strand synthesis.[1,22]

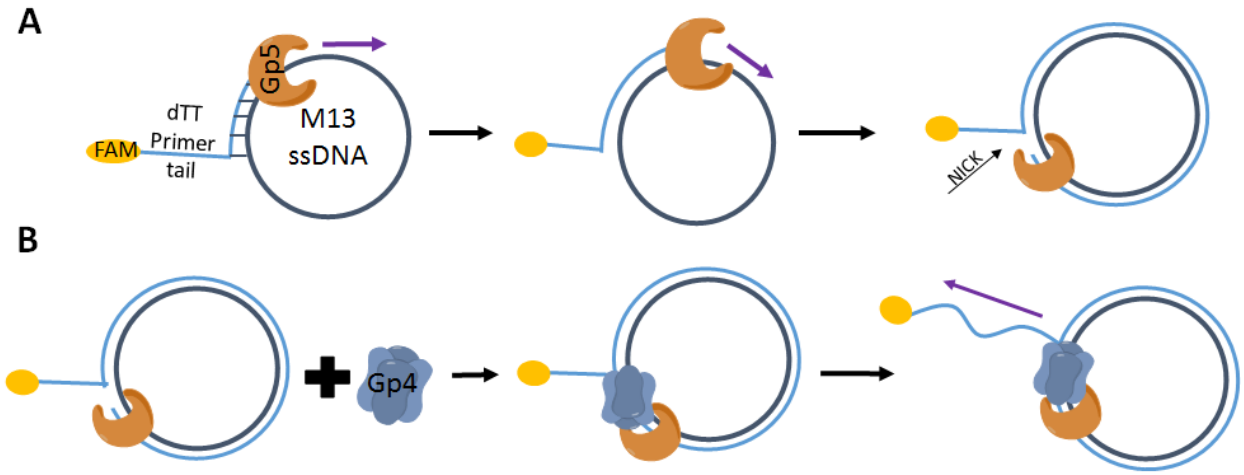

Figure 14. The rolling circle assay tests the replisomes ability to synthesize DNA. A) The ssM13 DNA is annealed to the FAM labeled primer. Without the helicase the wild type polymerase will be able to synthesize the circle piece of DNA until it encounters dsDNA. Without the helicase the wild type polymerase can not perform strand displacement synthesis and will stop once it synthesizes around the M13 once. B) When helicase is added to the reaction the polymerase and helicase can continually synthesis the circular DNA. Without the small RNA primers the complex cannot synthesize lagging strand and only the leading strand will extend.[gp5=polymerase; gp4=Helicase]

By adding the helicase we can determine if the mutated amino acids of the Fbp are vital for the electrostatic interaction between the helicase and the polymerase. Previously the entire Fbp was neutralized and it was shown that this interaction can no longer occur and the polymerase can not synthesize DNA on a dsM13 template.[2] We hypothesize that mutating any one of the amino acids that displayed a phenotype in the phage complementation assay they will have diminished electrostatic interaction and pol will only be able to replicate the ssM13 to the double stranded product.

2.4 SAXS

Small angle X-ray scattering (SAXS) is a method used to characterize the structure and shape of molecules in solution. It is also used to calculate the molecular weight and the concentration of homogeneous molecules in solution making it possible to determine a molar ratio.[23] The T7 replisome consisting of T7 polymerase, Trx, and primase-helicase was crystallized and the X-ray structure was determined by Dr. Wallen previously. This crystal structure shows three polymerases bound to a hexameric helicase, pol A, B, and C (Figure 15).

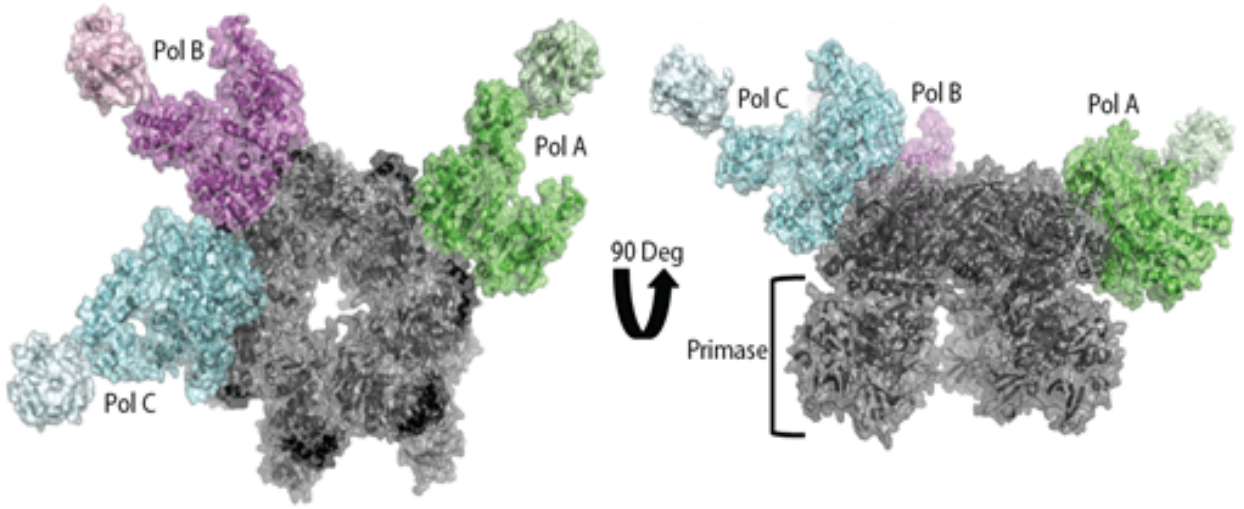

Figure 15. Crystal structure showing three polymerases bound (A, B, and C) to hexameric pirmase-helicase. Polymerase A and B are bound to the helicase in the same orientation while polymerase C is rotated 90° and leaning toward the outside of the helicase ring. Polymerase A and B are hypothesized to be the polymerases that are used to replicate both strands of DNA while pol C is bond to helicase in the event that either of the other pols dissociate from the complex.

In this crystal structure polymerases A and B are bound to helicase in the same orientation. Polymerase C is rotated about 90° and is leaning toward the outside of the helicase

ring. Only two copies of polymerase are needed to synthesize the leading and lagging strands DNA. This structure shows that polymerase A and polymerase B are positioned to replicate both strands while polymerase C might be bound in case pol B or C are dissociated. SAXS was performed on this complex to observe the structure in solution. The SAXS data was analyzed using FoXS.[24] This data was used to confirm that in solution this complex is the same as the crystal structure.

CHAPTER 3 EXPERIMENTAL

3.1 CLONING

All DNA samples were sequenced in-house using the Big Dye Terminator Cycle Sequencing Kit v.1.1 on the ABI 3730xl Sequencer (Applied Biosystem, USA) using 80cm capillary tubes for plasmid DNA or from Genewiz using T7, T7 term, and GP1-forward primers.

Mutations were introduced into the T7 gene 5 (polymerase) located in a pET21b plasmid by site directed mutagenesis.[25] Primers were obtained from Integrated DNA Technologies (IDT). The plasmid was then transformed into DH5α *E. coli* cells to amplify the mutated DNA produced. To confirm that DNA was made during the site directed mutagenesis the samples were ran on a 0.8% agarose gel. The samples containing DNA were then purified using a standard miniprep kit from Qiagen and the plasmid DNA was sequenced to confirm the correct mutation.

3.2 PROTEIN PURIFICATION

All columns for protein purification were purchased from either Biorad or GE Healthcare.

gp 2.5 Single Stranded Binding Protein: The pET21b vector that was used contains the full length SSB containing a precision protease site for the N-terminal His tag. This vector was transformed into BL21 (DE3) cells for overexpression of SSB. SSB was purified as previously described.[19] The SSB was first loaded onto a nickel affinity column. Equilibrated with buffer A: 50mM Tris pH 8, 200mM NaCl, 10mM imidazole, 10% glycerol, and 1mM BME. Buffer B containing 250mM imidazole was used to elute the protein with a gradient of 0-100% over 100mL. PreScission Protease was added to the purified protein at a ratio of 1:100 and the mixture was dialyzed overnight at 4 °C against 20mM Tris-HCL pH 7.5,

200mM NaCl, 0.1mM EDTA, 1mM DTT, and 5% glycerol. The protease was then removed by running the SSB over the Hi-Trap Nickel column again. The protein was further purified with a Hi-trap Q column followed by a Sephacryl HR 26/60 gel filtration column. Buffer A for the Q column contained 20mM Tris pH 7.5, 1mM EDTA, 1mM DTT and 5% glycerol. Buffer B for Q column is the same as A but contains 1M NaCl. Buffer B was used to elute the protein from Q column at a gradient of 0-100% B over 100mL. Buffer A for column Q was also used for the gel filtration column. The protein was identified using a 12% SDS-PAGE gel and pooled together and concentrator down using a concentrator. The extinction coefficient to determine concentration is 1.28g/L or 32890 M^{-1} Cm^{-1}. 1.7 mg/mL, or 20μM, was purified from 6 liters of *E. coli*.

gp5 Polymerase/Thioredoxin: The gp5 polymerase and Trx are expressed and harvested separately and mixed together in order to purify the gp5/Trx complex. HMS174(DE3) cells are used to co express both polymerase and RSF-Trx to limit proteolysis. Expression of Trx is low and therefore the Trx is separately expressed using the pTRX plasmid. To co-express polymerase in a pET21b plasmid and RSF-Trx, 6L of HMS174 (DE3) cells were grown at 37°C to 0.6-0.8 OD600 and the cells were induced with 1mM IPTG for 3 hours at 37°C. The cells were then harvested by spinning at 3500 RPM for 20 minutes and the bacterial pellet was re-suspended in 120mL of 25mM KPO4 pH 8.0, 600mM KCl, and 10% glycerol. This solution was then stored at -80°C.

Thioredoxin is over expressed in HMS174 (DE3) cells using the plasmid pTRX. 6L of cells were grown at 37°C for approximately 6 hours to 1.5 OD600 and then induced with 0.5mM IPTG and left to grow overnight at 37°C. The cells were harvested by spinning at 3500RPM for 20 minutes and the pellets were washed with 50mM Tris-HCl pH 7.5, 20mM EDTA, and 10% glycerol. Cells were then spun at 3500RPM for 20 minutes and the bacterial pellets were resuspended in 80mL of 50mM Tris-HCl pH 7.5, 1mM EDTA, and 10% glycerol

and stored at -80°C.

Once both proteins were available, both the Trx cells and pol/Trx cells were thawed and mixed together. Not only does the polymerase become more active by mixing thioredoxin with the polymerase but there is also a lesser chance of proteolysis. PMSF (0.2mg/mL) was added to the suspension and the mixture was left to stir at 4°C for 30 minutes. The cells were then lysed by sonication and spun down at 17000 RPM for 50 minutes to pellet the cell debris. The supernatant was collected and the nucleic acids were precipitated by slowly adding 0.5% final concentration of polyethyleneimine (PEI) with stirring at 4°C for 20 minutes. The suspension was spun at 17000 RPM for 40 minutes to pellet the DNA. The supernatant was collected and 60% saturation of ammonium sulfate was added to salt out the protein. It was stirred for 60 minutes and spun down at 17000 RPM for 20 minutes. The supernatant was then removed and the protein pellet was resuspended in 100mLs of 20mM Tris-HCl pH 7.5, 1mM EDTA, 1mM DTT, and 5% glycerol.

The resuspension was loaded on an ion exchange column, (HiTrap Heparin) HP column, that was equilibrated with Buffer A (20mM Tris-HCl pH 7.5, 1mM EDTA, 1mM DTT, and 5% glycerol). The protein column was washed with 30mL of buffer A and the protein was eluted over a gradient of $0-100\%$ Buffer B(1M NaCl, 20mM Tris-HCl pH 7.5, 0.1mM EDTA, 1mM DTT, 5% glycerol) while collecting 3mL fractions. Identification of protein was done with SDS-PAGE and the polymerase fractions were pooled together. The protein was diluted to a 100mM salt concentration with buffer A and loaded onto Hi-Trap Q HP column preequilibrated with Buffer A. Protein was eluted with a gradient of $0-80\%$ NaCl. The protein was collected in 3mL fractions and identified using SDS-PAGE. The fractions were pooled and dialyzed for 3 hours in dialysis buffer (10mM KPO_4 pH 7.4, 5% glycerol and 1mM BME). The dialyzed protein was then loaded onto Mini CHT Ceramic Hydroxyapatite preequilibrated with Buffer A (10mM KPO_4, 5% glycerol, and 1mM BME). The column was then washed with 30mL of 100% A and the protein was eluted with a gradient of $10-500$mM

KPO_4 collecting 1mL fractions. The protein was identified by SDS-PAGE and the fractions were pooled together and concentrated down to 1mL to be injected into a GE Sephacryl HR 26/60 gel filtration column. The protein was eluted with 20mM Tris-HCl pH 7.5, 100mM NaCl, 0.5mM EDTA, 1mM DTT, 5% glycerol. Pol/Trx complex is about 92KDa and gel filtration shows that the protein elutes as a monomer. The fractions are pooled and identified using a 15% SDS-PAGE gel. The sample was concentrated down using a concentrator. The extinction coefficient to determine concentration is 1.644g/L or $150230 M^{-1}\ cm^{-1}$. 3.5mg/mL or 13μM, was purified from 6 liters of *E. coli*. The protein was aliquoted and flash frozen in liquid nitrogen to be stored at -80°C.

The R590A mutant was successful expressed and purified as described for WT pol. The R590A was concentrated down to 11μM or 0.5mg/mL.

3.3 HARVESTING T7 PHAGE

To study the effects of the mutations of the T7 DNA polymerase we need to grow and harvest Δgp5 phage.[26,27] 0.05-0.2mL of Δgp5 phage were mixed with 0.25mL BL21 *E. coli* and 2.5 mL top T-broth agar(0.1% tryptone, 0.005% NaCL, and 0.7% agar). This mixture was poured evenly on the T-broth bottom agar (1% agar). The plates were incubated overnight at room temperature or at 37°C for 2-3 hours. A single plaque was picked and added to 5mL fresh overnight *E. coli* culture and 15mL T-broth in a 125mL Erlenmeyer flask. The culture was shaken at 30°C for 2-4 hours or until lysis occurs. 1g of NaCl is added to the culture and then spun at 10,000 RPM for 10 minutes. The supernatant was collected in a sterile tube and filtered using a 0.22μm bottle top syringe. The virus was stored at 4°C. From 15mL T-broth we obtained a viral titer of 7.8^9 PFU/mL.

3.4 PHAGE COMPLEMENTATION

Phage complementation was used to determine the effect the specific mutations have on viral replication in vivo.[28] BL21 *E. coil* cells with either the WT or mutant gp5 were grown overnight in 5mL of lysogeny broth (LB). Ampicillin is included in broth and agar to insure plasmid remains in the cells. The next morning 100μL of the overnight culture was added to 3mL of LB and allowed to grow to an OD600 of $\sim$1 at 37 °C. 300μL of the culture was placed into a 1.5mL eppendorf tube and mixed with 200μL Δgp5 T7 virus at a concentration of 1x10^{-8}PFU/mL. Top agar(0.7% LB agar) was melted and kept at 55 °C and 1mL with 100μ mL ampicillin was added to the bacteria viral solution. The solution was mixed by pipetting up and down several times and then plated onto 1% LB bottom agar using three well plate. The plates were incubated for 4 hours at 37 °C and the viral plaques were counted to determine a titer or plaque forming units per milliliter by the equation

$$\frac{PFU}{mL} = \frac{n}{VD} \tag{3.4.1}$$

where n is defined as the number of plaques counted, V is amount of viral sample plated, and D is the dilution factor of the virus.

3.4.1 SPOT TEST

Spot tests were also performed to determine the viral titer range to add to the plaque assays. This was done by plating the bacteria (grown to an OD 0.6-0.8) and LB top agar on to LB bottom agar plates. The top agar was left to harden and dry under a sterile hood for 5 minutes and the virus was serially diluted 10-fold for a range of $11x10^{-2}$ to $1x10^{-8}$, then 2μL was spotted onto the bacterial lawn. If the mutated phage is able to grow in vivo plaques will form where the virus was spotted. Viral plaques appear to be small holes in the bacterial lawn.

3.4.2 GROWTH CURVE

The growth curve assay determined the effect the mutations had on the growth rate of the phage. BL21 *E. coli* cells, with the pET21 plasmid containing polymerase, were inoculated in 5mL LB with 100μg/mL ampicillin and let to grow overnight at 37°C. 1mL cultures containing 100μg/mL of ampicillin were inoculated with 100μL of overnight culture are grown to OD 0.6-0.8. They were infected with Δgp5 T7 phage at an MOI of $1E^{-9}$. 50μL samples were taken at time 0 and every 2 hours for 6 hours and one sample was taken overnight at 15 hours. These samples were spun down to remove all cells and the supernatant was collected and stored in the fridge. Once all samples were collected each supernatant sample was serially diluted 10 fold 7 times.

Wild type *E. coli* cultures were grown to an OD 0.6-0.8 at 37°C. The *E. coli* cells were mixed with top agar heated to 50°C, at 200μL cells/mL top agar. The top LB agar mixture was plated on top of LB bottom agar also containing ampicillin. Once the top agar solidified 5μL of collected samples were spotted on the top agar for each time point and each dilution. The plates were let to incubate at 37°C for 4 hours. At 4 hours the individual plaques were counted for the lowest dilution of viral growth to obtain a viral titer. Wild type polymerase was compared to four mutations (R590A, F487A, I569A, and R687A).

3.5 ROLLING CIRCLE

To study the polymerases ability to interact with the helicase and synthesize DNA rolling circle replication assay can be used. We can look at the effects the single mutations of the Fbp and the exonuclease site have on the pol ability to synthesize DNA.

Sample (45μL) samples containing 500nM dNTP's,20nM primed ssM13, and 20nM polymerase were prepared in 25mM tris-HCl pH 7.5, 100mM NaCl, 5mM MgCl2, and 5mM DTT buffer. Samples before adding enzyme were preheated at 37°C in a water bath. The

reactions were started by the addition of enzyme, and all assays were performed at 37°C. Samples were quenched at 10s, 30s and 2min with 50mM EDTA, 0.2% SDS and 10% glycerol. Samples were incubated at 37°C until quenched. The samples were loaded onto 0.6% agrose gel containing ethidium bromide dye (EtBr) and ran at 150V for 1.5 hours. The gel was imaged using a Biorad ChemiDoc MP imager.

Table 1. M13 primer used for this study

DNA Oligos	Oligo Sequence
M13_rollingprimerFAM	5'FAM-(dT35)GTGTAAAACGACGGCCAGTGC

CHAPTER 4 RESULTS AND DISCUSSION

4.1 PURIFICATION

T7 DNA polymerase had been purified previously using phosphocellulose as the first column.[29] This column is no longer available on the market, and therefore part of this thesis was focused on developing a new method of purification. Heparin was found to be a suitable replacement for Phosphocellulose, as the polymerase bound and eluted well from the column. The columns are in order HiTrap Heparin HP, Hi-Trap Q, hydroxyapatite, and sepharyl HR 26/60 gel filtration. WT DNA polymerase and R590A were successfully purified using this method. WT polymerase was purified at 3.5mg/mL while R590A was purified at 1.19mg/mL.

WT polymerase was first purified using Heprin, a cation exchange column. Figure 16 shows the chromatogram and the SDS gel of the fractions for the first column.

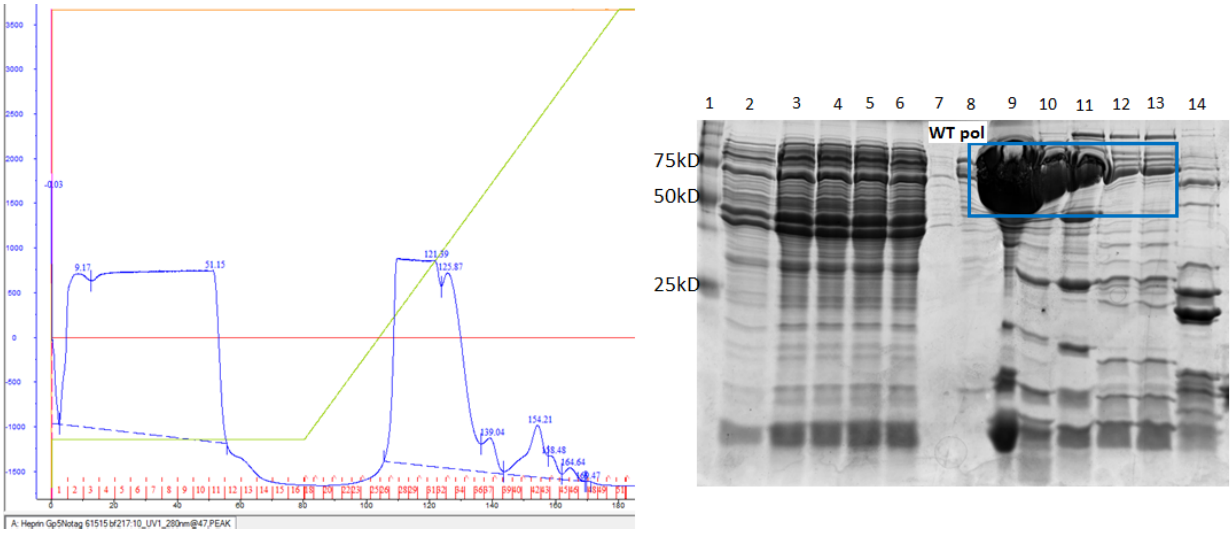

Figure 16. The chromatogram on the left after the Heprin column. Fractions 1-12 are flow through represented on the gel in lanes 2-7. Lanes 8-13 contain every other fractions from 26-36 which contain WT pol. Fractions 27-33 were pooled to eliminate impurities. These fractions were diluted to 10% salt and loaded onto Hi-Trap Q column.

After running WT pol over the Heparin column fractions 27-33 were pooled together and diluted to 10% salt. The pol was then loaded onto a Hi-Trap Q ion exchange column. Figure 17 shows the chromatogram and SDS gel of the fractions for the Q column.

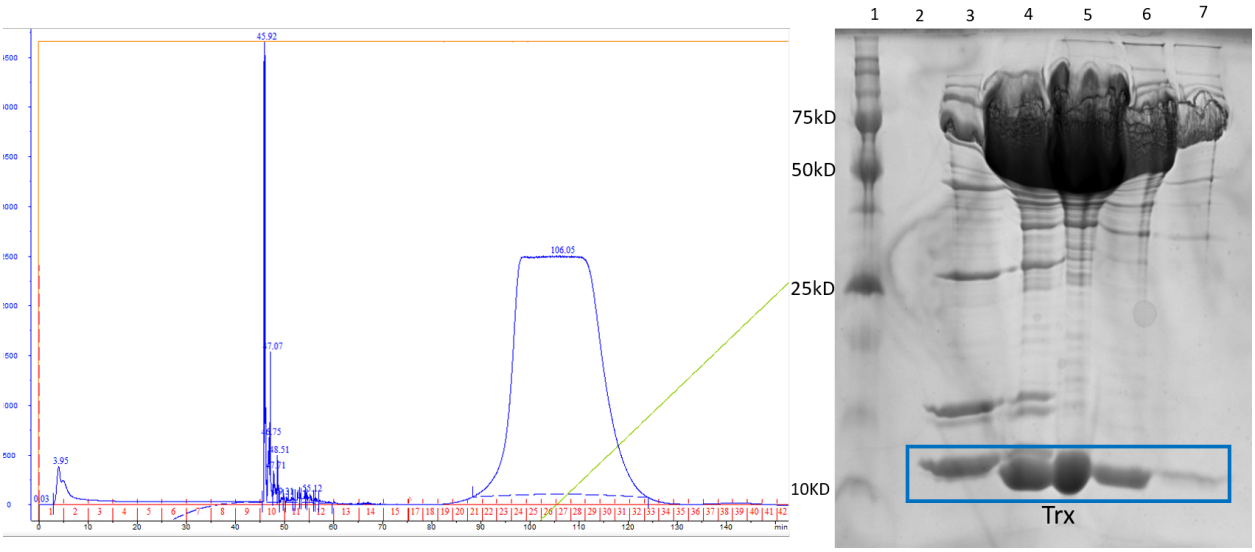

Figure 17. The chromatogram on the left is after running the pol over the Q column. Fractions 10-12 contain a bubble. Pol and Trx co-purify and Trx is marked by a blue box in the SDS gel. Fractions 20-33 contain the WT pol/Trx. Every other fraction was collected from 23-31 and are on lanes 3-7 on the SDS gel. These lanes contain WT pol/Trx. All fractions 23-31 were pooled and put on dialysis buffer for 3 hours before being loaded onto the hydroxyapatite column.

Fractions 23-31 were pooled after the Hi-trap Q column. These fractions containing the protein were put on dialysis buffer for 3 hours before being loaded onto the Hydroxyapatite column. Figure 18 shows the SDS gel and chromatogram for the Hydroxyapatite column. The Hydroxyapatite column is used to purify the polymerase fully bound to Trx to insure that the polymerase is active. This column is an ion exchange column and mimics the DNA backbone, making it an ideal column to purify the active T7 DNA pol/Trx complex.

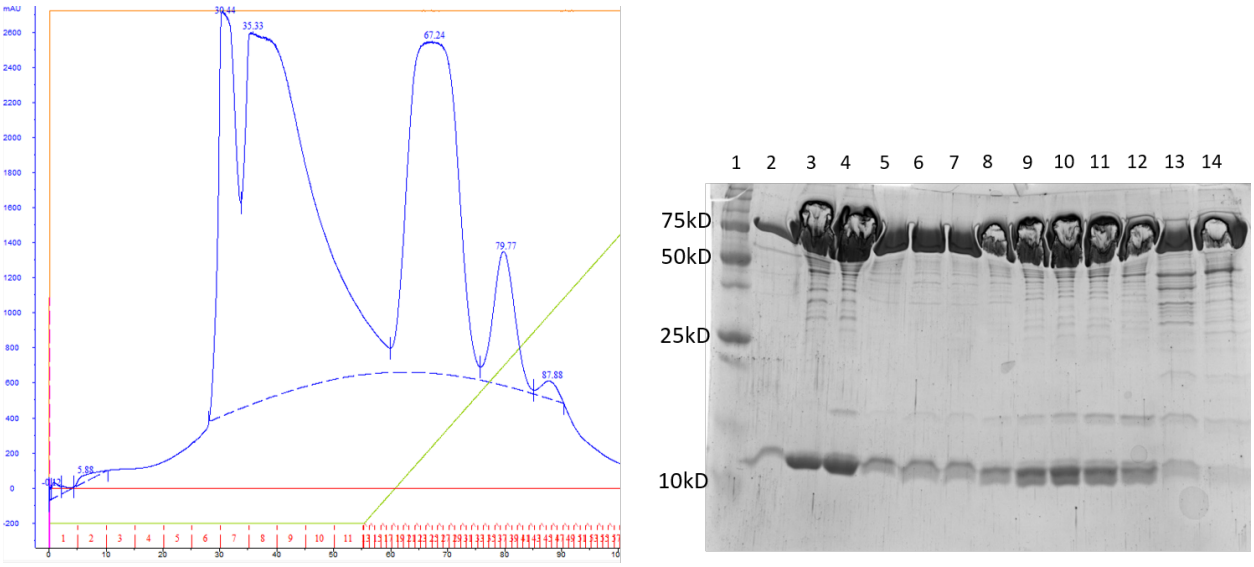

Figure 18. The left panel shows the chromatogram for pol/Trx complex after being eluted off of the Hydroxyapatite column. The Hydroxyapatite column is an ion exchange column and mimics the DNA backbone. This column is used to purify the most active polymerase bound to Trx. The chromatogram shows 5 peaks, the first two are from being overloaded as seen with littler separation of the two peaks. Peaks 3 and four contain are well separated and contain pol/Trx. The third peak contains the most active pol/Trx complex while the fourth peak is less active. The right panel is of the SDS gel for the eluted sample. Lanes 2-5 contain fractions from the first and second peak. Lanes 7-12 contain every other fractions 19-30 and lanes 13-14 contain fraction 39 and 45. The SDS gel shows that peak 3 contains both polymerase and high concentration of Trx indicating this is the most active polymerase. Fractions 19-30 were pooled and concentrated down to 1mL and loaded onto the sizing column.

Once the protein was eluted fractions containing the active pol/Trx (fraction 19-30) were pooled and concentrated down to 1mL. The concentrated protein was loaded onto the sizing column. Figure 19 shows the chromatogram and SDS gel of eluted protein.

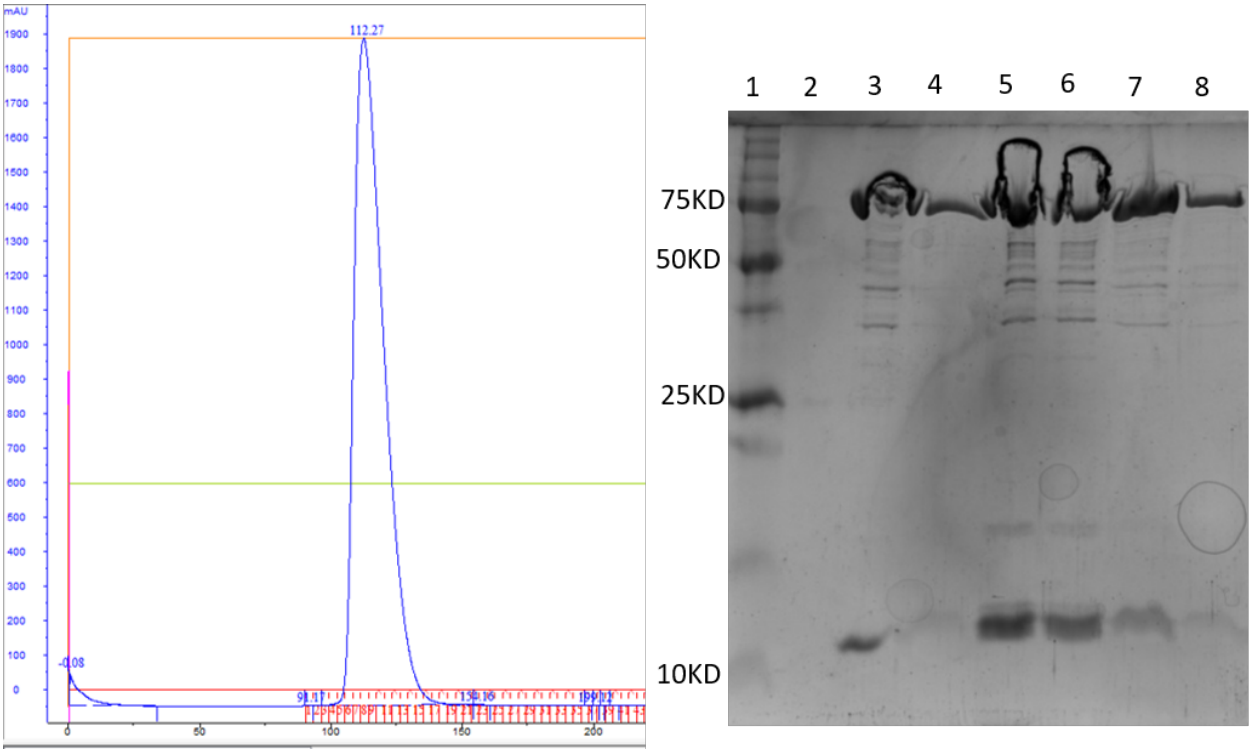

Figure 19. Chromatogram on the left shows the pol/Trx started to elute at 100 minutes from fractions 5-17. The SDS gel on the right compares pure WT pol/Trx in lane 3 to the fractions sampled (5-17). T7 DNA Polymerase has a molecular mass of $\sim$ 80Kda and it binds to Trx with a mass of $\sim$ 12kda.[1] According to the SDS gel we have successfully purified pol/Trx complex.

Fractions 5-17 were pooled and concentrated down after elution from sizing column. The protein was sucessfully purified usin this new method. The protein was concentrated to 3.5mg/mL for WT and 1.7mg/mL for R590A.

4.2 PHAGE COMPLEMENTATION

The Δgp5 phage was complemented with a plasmid transformed into BL21 *E. coli* cells. Phage complementation tests how well the phage can survive in vivo when mutations are introduced into the essential DNA polymerase. Amino acids of the Fbp, hydrophobic pocket, and exonuclease of the polymerase were mutated and tested using phage complementation.

4.2.1 FRONT BASIC PATCH

To determine the appropriate concentration of virus for a plaque assay, a spot test was performed. The mutations of the Fbp and hydrophobic pocket are compared in Figure 20. WT polymerase is seen to form plaques at $1x10^{-8}$ dilution while the mutations F487A, R687A, and I569A form plaques only to $1x10^{-5}$ dilution. R590A being the weakest of all the mutations only forms plaques to the $1x10^{-4}$ dilution.

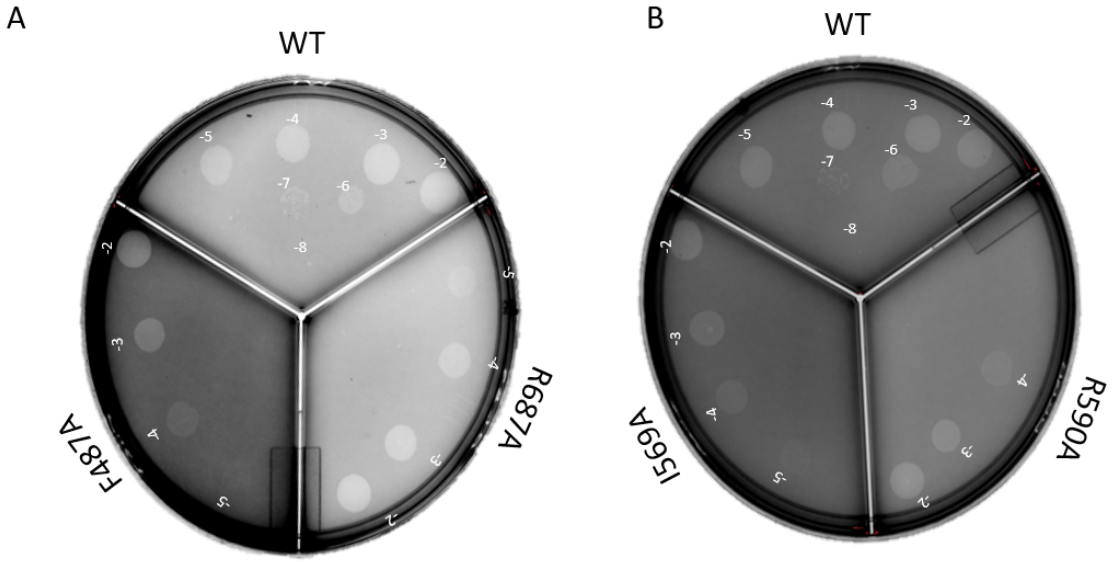

Figure 20. The spot test performed showed wild type pol forming plaques at $1x10^{-8}$ while F487A, R687A, and I569A could only form plaques at $1x10^{-5}$ dilution. R590A was even more of a detrimental mutation only being able to form plaques at $1x10^{-4}$ dilution. This assay demonstrates how the point mutations have impaired the phage survival.

When performing a plaque assay the amount of virus that is added is very important. If you add a concentrated amount of virus the effects of the mutations may not be seen. Since WT can form plaques at the lowest dilution ($1x10^{-8}$) the plaque assays will be plated at this dilution. Table 2 show the titer calculated for each mutation for the plaque assays performed. WT gives a titer of $2.8x10^{9}$ while R590A, F487A, and I569A did not form plaques

at the same dilution. The results suggest that these amino acids play an essential role in replication. R687A had a different phenotype than wild type, as it did form plaques but they were much smaller making them hard to count. The smaller plaques suggest that the virus is sick as compared to WT, which suggests that R687 does play an important role in replication.

Table 2. Titer of Δ gp5 T7 phage plated onto *E. coli* containing WT pol. The phage was plated at a 10^{-8} dilution.

Pol	pfu/mL at 10^{-8} dilution
WT	$2.8x10^{9}$
R590A	no plaques observed
F487A	no plaques observed
I569A	no plaques observed
R687A	plaques visible but very few and too small to count

4.2.2 EXONUCLEASE

To understand the role of W160 in exonuclease activity, a spot test was also performed on the exonuclease mutations (Figure 21). For this assay WT formed plaques to $1x10^{-7}$ dilution. Seen in Figure 21, the behavior of the mutations are identical to WT and form plaques to $1x10^{-7}$ dilution. There are several possibilities 1) that error prone synthesis is occurring and the phage is still able to survive 2) there are other proteins in the host that help repair DNA damage caused by mistakes made by the T7 DNA polymerase. Figure **??** demonstrates that 5A7A and W160A are unable to fix mistakes on small primer/template substrates.

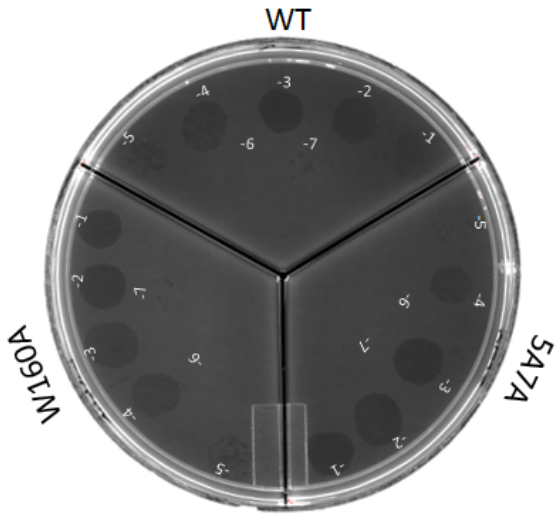

Figure 21. The exonuclease mutations were not seen to be detrimental to the phage and both W160A and 5A7A were able to form plaques at the same dilution as WT, $1x10^{-7}$.

The countable spots at the lowest dilution were determined for each polymerase. Table 3 is the calculated titer for WT, W160A, and 5A7A. The titers found for the mutations are similar to WT.

Table 3. Titer of WT, 5A7A, and W160A from spot test. All three were seen to form plaques at $1x10^{-7}$ dilution

Pol	pfu/mL
WT	$2.4x10^{12}$
W160A	$1.8x10^{12}$
5A7A	$2.6x10^{12}$

4.3 GROWTH CURVE

To study the effects of the mutation on growth of the phage over time, a multi-step growth curve was performed. BL21 *E. coli* cells were infected with a $1x10^{-7}$ MOI of phage. These cells were grown for 15 hours and samples were taken every two hours for 6 hours and 1 sample at 15 hours. At this MOI, WT pol was seen to plateau at 4 hours at a titer of

1.4x10^7 PFU/mL while R590A did not show any growth (Figure 22). These results show that at an MOI that supports WT phage growth, R590A is unable to replicate the phage DNA.

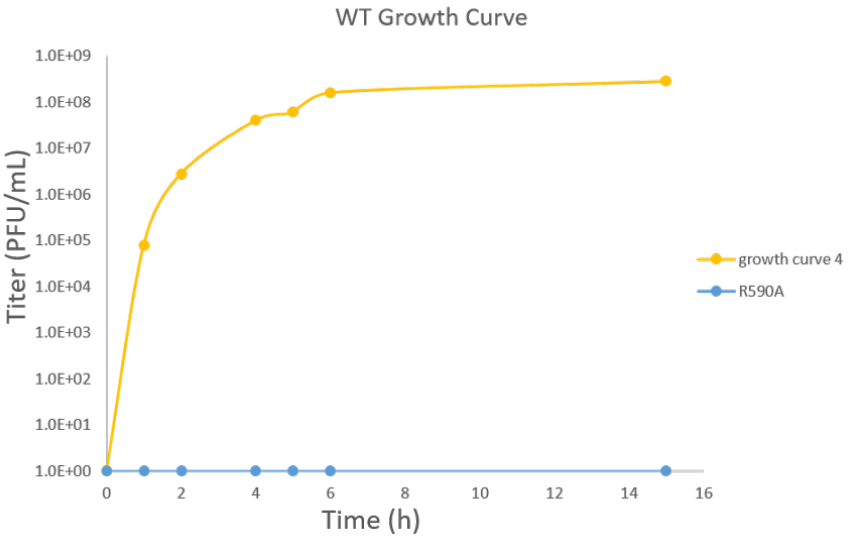

Figure 22. Growth curve of WT and R590A at an MOI of 1x10^{-7}. These samples were grown for15 hours. R590A shows no growth for 15 hours while WT grows to a titer of 1.4x10^7 PFU/mL before leveling off. R590 is a vital amino acid and when mutated to an alanine can no longer grow in vivo.

Preliminary data for the F487A and I569A mutants are seen in Figure 23. To see the effects of these mutations the MOI was much lower than that of R590A. Samples were grown for 15 hours and samples were taken every hour for 5 hours and one sample at 15 hours. As observed perviously, WT phage grows and plateaus at 4 hours at a titer of 1.4x10^7 PFU/mL at an MOI of 1x10^{-9} WT plateaus at 5 hours at titer of 6x10^{12}. At this MOI the WT phage can reach a higher concentration of virus before plateauing. The mutants F487A and I569A both plateau at 4 hours with a titer of 3.6x10^{10} PFU/mL and 1.2x10^{10} PFU/mL respectively. This data was not completed in triplicate and needs to be continued in future work. Although, these mutants are not as detrimental as R590A, they still impact the phages ability to replicate. This data suggests that the electrostatic interactions (R590) with F566

appears to be more important than the hydrophobic pocket itself. This however is limiting, as we need to perform more experiments.

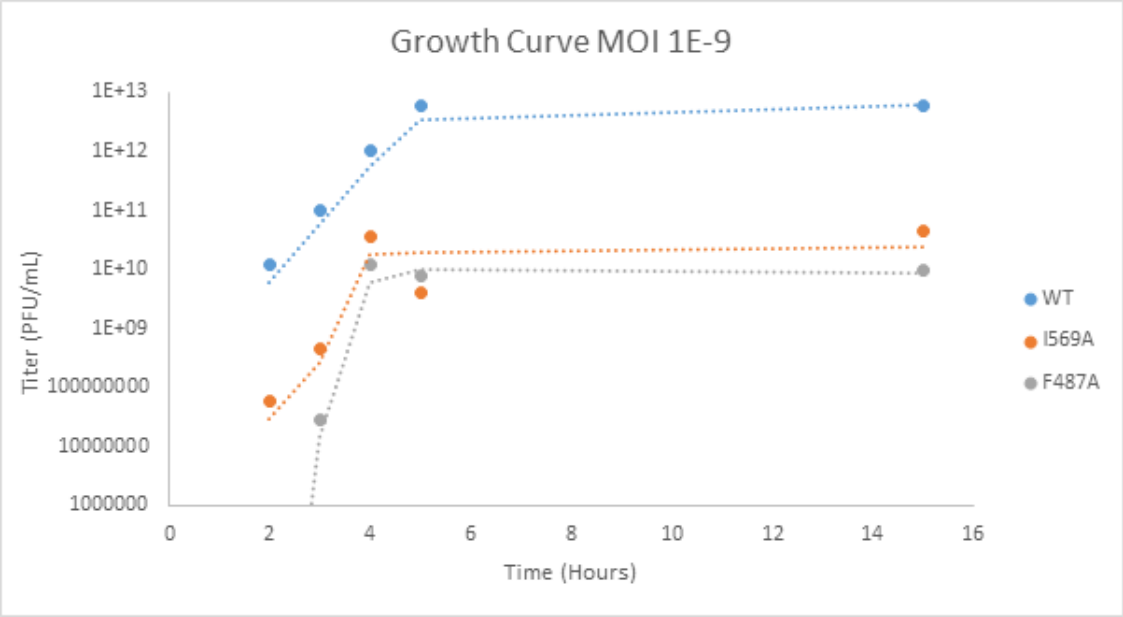

Figure 23. Preliminary data shows the mutants are impaired, in their growth over time compared to wild type, at an MOI of $1x10^{-9}$.

4.4 ROLLING CIRCLE

We recently optimized the rolling circle assay. Several kinds of gels were attempted to obtain sharp and visible bands. We were able to optimize an acrylamide gel with EtBr for this assay. The ssM13 circular DNA is annealed to the M13_rollingprimerFAM (Table 1). Since the primer has a FAM label we wanted to image the gel without any dye but the imager at Western Carolina University is not sensitive enough to read 50nM FAM labeled DNA. Rolling circle demonstrates the ability of the pol to synthesize DNA and is able to test all the mutants created. The ssM13 DNA is 7,249bp in length and is single stranded. If the polymerase can synthesize DNA we hypothesize that it will synthesize the double stranded M13 product one time. We hypothesized that the mutations of the Fbp would not

effect the ability of the polymerase to synthesize DNA since it is not related to the active site. However, due to results from previous experiments of the exonuclease mutants 5A7A and W160A, we hypothesized that the polymerase would not be able to complete the ssM13 primer template. Since this site allows the polymerase to fix mistakes made if it is no longer active previous data has shown that the polymerase will lock onto a linear primer template and can no longer continue synthesis (Figure10). Future work will use this assay to look at the electrostatic interaction between the pol and C-terminal tail of the helicase. As of now we are only able to test if the mutants are able to synthesize DNA.

For this assay WT, 5A7A, R590A/5A7A, and W160A were used. Figure 24 shows the acrylamide gel for each reaction. The reactions were quenched at 10s, 30s, 2min, and 60min. All polymerases are able to synthesize the dsM13 product at ∼2min. At 60min WT polymerase and W160A have completed the dsM13 product as hypothesized. What is interesting is 5A7A and R590A/5A7A seem to have the ability to perform strand displacement which disagrees with our hypothesis. If the polymerase loses its exonuclease activity then the polymerase is able to unwind dsDNA and continue on in the rolling circle.

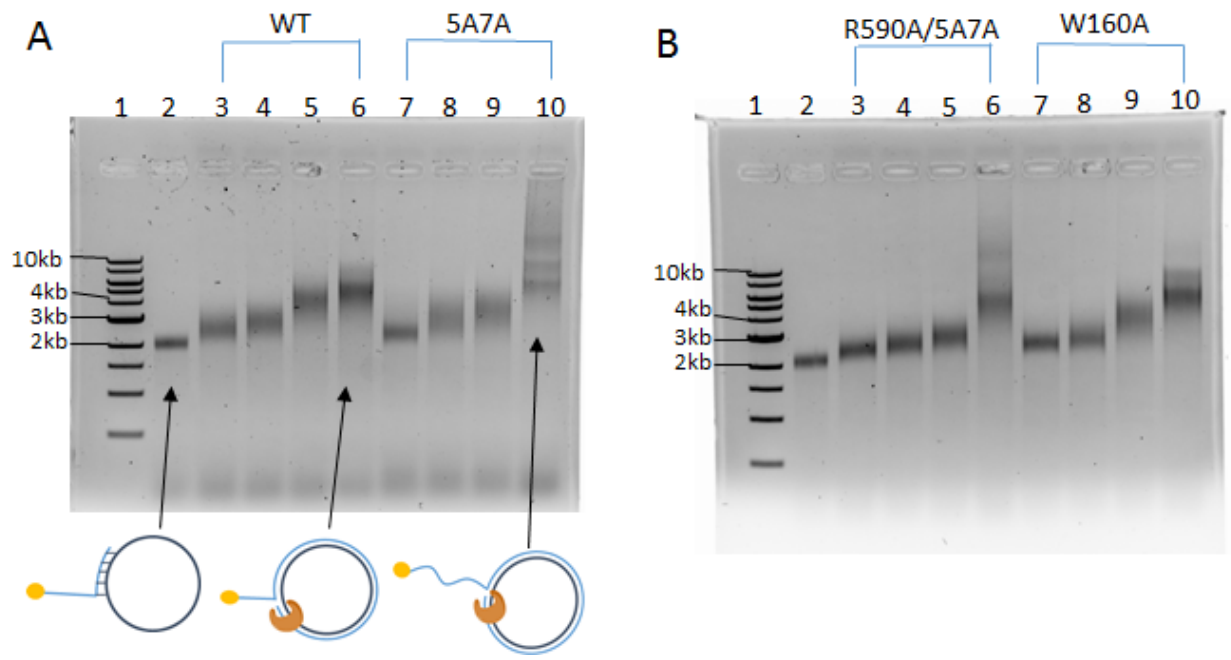

Figure 24. Rolling circle results. For each reaction, aliqupts were quenched at 10s (lanes 3 and 7), 30s (lanes 4 and 8), 2min (lanes 5 and 9), and 60min (lanes 6 and 10). A) WT polymerase and 5A7A mutant. Lane 2 contains M13 DNA only with the FAM labeled primer. At 2min (lane 5) the WT polymerase has completed synthesizing the M13 DNA and it is now double stranded. 5A7A continues synthesizing the DNA and is seen to perform strand displacement synthesis for lane 10 at 60s. The mutation in the exonuclease site allows the polymerase to continue on. B) Lane 2 contains ssM13 DNA and FAM labeled primer. R590A/5A7A is also seen to perform strand displacement synthesis lane 6 at 60s. W160 is similar to WT and it completes the dsM13 product at 2min and is completely finished at 60min.

Future work will be to further test rolling circle by adding helicase or a mismatch to the primer template. By adding helicase we will be able to see how well the Fbp and hydrophobic pocket mutants can perform strand displacement. If a mismatch is added to the primer template we will be able to see how the exonuclease mutants can synthesize DNA faced with a known mutation that needs to be corrected.

4.5 SAXS

The SAXS data collected was modeled using the program FoX'S doc. The data for the structure in solution was hypothesized to be similar to that of the crystal structure (Figure 25). However, the model data of the polymerase B bound to a hexameric helicase is different. When comparing the data generated from FoX's doc the best fit when comparing to the crystal structure is a model of the DNA polymerase lying across the C-terminal face of the helicase. Compared to the crystal structure this model exposes the Fbp for what seems to be a way to stabilize the interaction with the C-terminal tail of the helicase. Depending on what is needed during replication the DNA polymerase is loosely bound to the Fbp so it can move around.

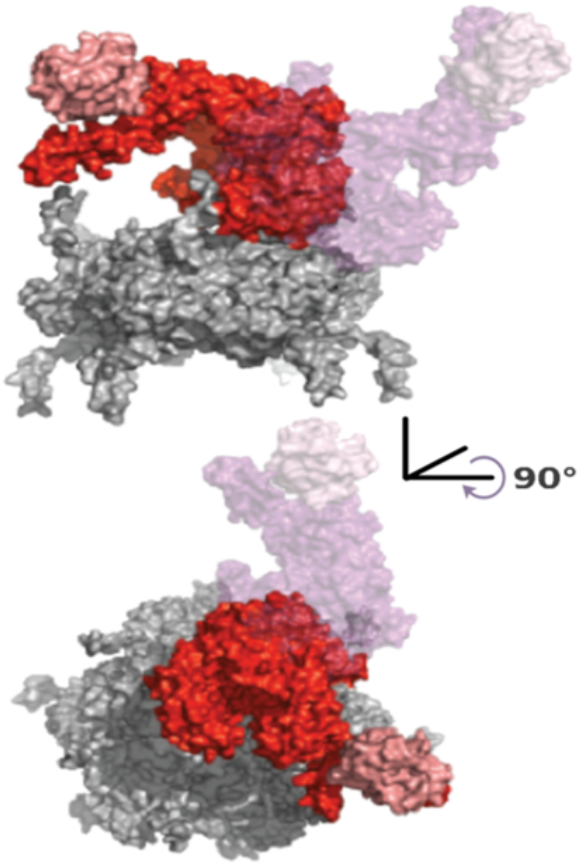

Figure 25. The crystal structure (magenta) reveals that the structure in solution differs from a model based on the interaction of Pol B with gp4 in the crystallized replisome. The polymerase in the model is leaning in toward the helicase ring. This model displays how the replisome has different modes of binding and is dynamic depending on what is needed for replication.

CHAPTER 5 CONCLUSIONS AND FUTURE WORK

This thesis focuses validating protein/protein interactions observed in a crystal structure. This crystal structure shows the essential C-terminal tail binding to the polymerase Fbp and hydrophobic pocket. The T7 replisome is a coordinated system, if one or more proteins cannot interact then DNA synthesis cannot occur. The aromatic nature of the C-terminal F566 is essential for DNA synthesis and if it is mutated or removed DNA synthesis cannot occur. It is important to understand how this tail interacts with the polymerase, as it can give insight into how more complex replication systems interact. This crystal structure also shows W160 in the exonuclease site stacking with the base of a nucleotide triphosphate. We hypothesize that this amino acid is used in stabilizing the DNA in the exonuclease active site.

To verify the importance of these interactions site directed mutagenesis was performed on the amino acids involved. In the hydrophobic pocket F487, I569, and L613 were mutated to an A. Two amino acids in the Fbp (R687, and R590) that form a salt bridge with the C-terminal tail were also mutated to A. In the exonuclease site the W160 was mutated to an A to determine if the aromatic nature of the amino acid is essential for stacking with the base of the nucleotide. Several assays were performed to verify these interactions including: phage complementation, growth curve assay, rolling circle replication assay, and primer extension assay. These assays helped to determine which of these amino acids, if not all, are essential for coordinating the replisome and DNA synthesis.

Phage Complementation: The spot test used for phage complementation demonstrated how the mutations effected the phage growth in vivo. WT polymerase was able to form plaques at the highest dilution of $1x10^{-8}$ while the mutants F487A, I569A, and R687A could only form plaques at $1x10^{-5}$ dilution. R590A formed plaques at an even lower dilution of $1x10^{-4}$. These mutations have caused a 3 log and 4 log difference in plaque formation

compared to WT. When plating these mutations and WT at a $1x10^{-8}$ PFU/mL a titer was calculated for each. WT had a titer of $2.8x10^{9}$ while no plaques were observed for R590A, F487A, and I569A. Plaques were visible for R687A but the plaque size was much smaller than WT and a titer could not be calculated. The smaller plaques suggest this mutation has weakened the growth of the phage showing that R687 is a vital amino acid. From the spot test and plaque assay we can conclude that these amino acids play an important role phage growth and survival.

The exonuclease mutants (W160A and 5A7A) were also tested using phage complementation. These mutations formed plaques at the same dilution (1×10^{7}), as WT. The titers were calculated at $2.4x10^{12}$, $1.8x10^{12}$, and $2.6x10^{12}$ for WT, W160A, and 5A7A, respectively. Since these mutations do not alter the growth of the phage we hypothesize that the polymerase can still survive even though mistakes made during replication are not corrected. We suspect the polymerase with these mutations make mistakes when synthesizing DNA since the exonuclease site is not active and mutations cannot be corrected. The phage could still survive if there are other proteins recruited to fix the DNA damage, or if the mutations are not detrimental to the phage.

Growth Curve: The amino acid R590 proved to be essential for replication in the multi-step growth curve assay. At an MOI of $1x10^{-7}$ WT grew for 4 hours before plateauing at a concentration of $1.4x10^{7}$PFU/mL while R590A did not show any growth over the 15 hour time period. These results show that, at an MOI that supports WT growth, R590A is unable to replicate phage DNA. We conclude that the interaction observed between R590 and the primase-helicase F566 main chain carboxyl is needed for efficient replication of the phage.

A lower MOI ($1x10^{-9}$) was required to see the effects of mutations F487A and I569A have on phage in *E. coli*. Preliminary data reveals that the WT grows for 5 hours until reaching the maximum concentration of phage at $6x10^{12}$PFU/mL. F487A and I569A grow

for four hours before plateauing at titers of 3.6^{10}PFU/mL and $1.2x10^{10}$PFU/mL. This data was not completed in triplicate and this study will be continued in the future. WHile these hydrophobic pocket mutations affected replication they were not as detrimental as the Fbp mutation R590A. This data suggests that the electrostatic interaction between the F566 and R590 appears to be of more essential than the hydrophobic pocket for replication. Completing this experiment in triplicate and determine the growth curve for R687A, W160A, and 5A7A will need to be continued.

Rolling Circle: Using ssM13 circular DNA we are able to investigate the ability of WT and our mutant polymerases to replicate DNA alone and its ability to interact with helicase. The ssM13 is annealed to a primer with a 5'-dt35 tail. The primer is needed for polymerase to load onto DNA and the 5' tail allows the helicase to load at the fork junction. Adding just polymerase to the DNA reveals how the mutations alter the ability to bind to and synthesize DNA. WT and W160A are able to synthesize the dsM13 DNA product at 2min. The 5A7A and R590A5A7A complete the circle, then continue around the circle for strand displacement synthesis. The ability of the 5A7A mutant to perform strand-displacement synthesis was completely unexpected. Mutating the exonuclease previously killed its ability to perform DNA synthesis on a short DNA substrate containing a mismatch. The polymerase without an exonuclease site has a low fidelity and during synthesis of M13 DNA mistakes might have been made. For every $10^3 - 10^5$ nucleotides incorporated ~ 1 mistake is made. If a mistake was made we hypothesized that the W160A or 5A7A would not be able to repair the mistake and DNA synthesis would halt. Instead W160A was able to synthesize the dsM13 DNA product like WT and 5A7A was able to perform strand displacement synthesis and continue around the circular DNA replicating M13. It seems that without the exonuclease activity the polymerase either did not create any mutations while replicating ssM13, which is unlikely or the enzyme disregards any mutations and continues synthesizing the DNA. For future work we would like to study this further by introducing M13 primers that contain

a mismatch. By adding a primer with a known mismatch we believe that the W160A and 5A7A will not be able to continue synthesis and halt. For replication assays with the Fbp mutants, we would also like to pursue a competition assay by adding primase-helicase and titrating in the C-terminal tail peptide. The helicase will compete with the peptide to bind to the hydrophobic pocket and at a higher concentration of peptide we should see diminished interaction between the helicase and polymerase. If the helicase cannot interact with the polymerase DNA synthesis will not occur. If the mutations in the hydrophobic pocket and the Fbp are essential, then only WT should lose the ability to perform strand displacement synthesis in the presence of primase-helicase.

REFERENCES

[1] Hamdan, S. M. & Richardson, C. C. Motors, switches, and contacts in the replisome. *Annual review of biochemistry* **78**, 205–43 (2009). URL `http://www.ncbi.nlm.nih.gov/pubmed/19298182`.

[2] Zhang, H. *et al.* Helicase-DNA polymerase interaction is critical to initiate leading-strand DNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 9372–7 (2011).

[3] Robinson, A., J. Causer, R. & E. Dixon, N. Architecture and Conservation of the Bacterial DNA Replication Machinery, an Underexploited Drug Target. *Current Drug Targets* **13**, 352–372 (2012).

[4] MacNeill, S. Composition and dynamics of the eukaryotic replisome: A brief overview. In MacNeill, S. (ed.) *The Eukaryotic Replisome: a Guide to Protein Structure and Function*, vol. 62 of *Subcellular Biochemistry*, 1–17 (Springer Netherlands, 2012).

[5] Baker, T. A. & Bell, S. P. Polymerases and the Replisome : Machines within Machines. *cell press* **92**, 295–305 (1998).

[6] Duderstadt, K. E., Reyes-lamothe, R., Oijen, A. M. V. & Sherratt, D. J. Replication-Fork Dynamics 1–17 (2014).

[7] Curth, U. *et al.* Single-stranded-DNA-binding proteins from human mitochondria and Escherichia coli have analogous physicochemical properties. *European journal of biochemistry / FEBS* **221**, 435–443 (1994).

[8] Maréchal, A. & Zou, L. RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response. *Cell Research* **25**, 9–23 (2015).

[9] Manuscript, A. & Responses, S. Replication to DNA Damage and Stress Responses **208**, 267–273 (2011).

[10] Qimron, U., Tabor, S. & Richardson, C. C. New Details about Bacteriophage T7-Host Interactions. *Microbe* **5**, 117–122 (2010).

[11] Lee, S.-J. & Richardson, C. C. Choreography of bacteriophage T7 DNA replication. *Current opinion in chemical biology* **15**, 580–6 (2011).

[12] Lee, J., Chastain, P. D., Kusakabe, T., Griffith, J. D. & Richardson, C. C. Coordinated leading and lagging strand DNA synthesis on a minicircular template. *Molecular cell* **1**, 1001–10 (1998).

[13] Huber, H. E., Tabor, S. & Richardson, C. C. Escherichia coli thioredoxin stabilizes complexes of bacteriophage T7 DNA polymerase and primed templates. *Journal of Biological Chemistry* **262**, 16224–16232 (1987).

[14] Hamdan, S. M. *et al.* A unique loop in T7 DNA polymerase mediates the binding of helicase-primase, DNA binding protein, and processivity factor. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 5096–5101 (2005).

[15] Notarnicola, S. M., Mulcahy, H. L., Lee, J. & Richardson, C. C. The acidic carboxyl terminus of the bacteriophage T7 gene 4 helicase/primase interacts with T7 DNA polymerase. *The Journal of biological chemistry* **272**, 18425–18433 (1997).

[16] Lee, S.-J., Marintcheva, B., Hamdan, S. M. & Richardson, C. C. The C-terminal Residues of Bacteriophage T7 Gene 4 Helicase-Primase Coordinate Helicase and DNA Polymerase Activities. *Journal of Biological Chemistry* **281**, 25841–25849 (2006).

[17] Kong, D. & Richardson, C. C. Role of the acidic carboxyl-terminal domain of the

single-stranded DNA-binding protein of bacteriophage T7 in specific protein-protein interactions. *The Journal of biological chemistry* **273**, 6556–6564 (1998).

[18] Marintcheva, B. *et al.* Mutations in the gene 5 DNA polymerase of bacteriophage T7 suppress the dominant lethal phenotype of gene 2.5 ssDNA binding protein lacking the C-terminal phenylalanine. *Molecular microbiology* **72**, 869–80 (2009).

[19] Hollis, T., Stattel, J. M., Walther, D. S., Richardson, C. C. & Ellenberger, T. Structure of the gene 2.5 protein, a single-stranded DNA binding protein encoded by bacteriophage T7. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 9557–62 (2001).

[20] Rezende, L. F., Hollis, T., Ellenberger, T. & Richardson, C. C. Essential amino acid residues in the single-stranded DNA-binding protein of bacteriophage T7. Identification of the dimer interface. *The Journal of biological chemistry* **277**, 50643–50653 (2002).

[21] Ghosh, S., Marintcheva, B., Takahashi, M. & Richardson, C. C. C-terminal phenylalanine of bacteriophage T7 single-stranded DNA-binding protein is essential for strand displacement synthesis by T7 DNA polymerase at a nick in DNA. *Journal of Biological Chemistry* **284**, 30339–30349 (2009).

[22] Hamdan, S. M., Loparo, J. J., Takahashi, M., Richardson, C. C. & van Oijen, A. M. Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature* **457**, 336–9 (2009).

[23] Kikhney, A. G. & Svergun, D. I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Letters* **589**, 2570–2577 (2015).

[24] Schneidman-duhovny, D., Hammel, M. & Sali, A. FoXS : a web server for rapid computation and fitting of SAXS profiles **38**, 540–544 (2010).

[25] *Molecular cloning : a laboratory manual* (Cold Spring Harbor Laboratory Press, 2012).

[26] Biology, M. & Universit, L.-c. F. Bacteriophages **501** (2009).

[27] Lehninger, A., Nelson, D. L. & Cox, M. M. *Lehninger Principles of Biochemistry* (W. H. Freeman, 2013), sixtgh edition edn.

[28] Lee, S.-J., Chowdhury, K., Tabor, S. & Richardson, C. C. Rescue of bacteriophage T7 DNA polymerase of low processivity by suppressor mutations affecting gene 3 endonuclease. *Journal of virology* **83**, 8418–8427 (2009).

[29] Wallen, J. R., Majka, J. & Ellenberger, T. Discrete interactions between bacteriophage T7 primase-helicase and DNA polymerase drive the formation of a priming complex containing two copies of DNA polymerase. *Biochemistry* **52**, 4026–36 (2013).